LIMIT THEOREMS FOR COMPETITIVE DENSITY-DEPENDENT POPULATION PROCESSES

TODD L. PARSONS

1. Notation

- We write $\mathbb{N}_0$ for the set of non-negative integers.
- We use Landau asymptotic notation, where all asymptotics are with respect to N, defined below. We write $f(N) = O(g(N))$ and $f(N) = o(g(N))$ if
$$\limsup_{N\to\infty} \left|\frac{f(N)}{g(N)}\right| < \infty$$
and
$$\lim_{N\to\infty} \left|\frac{f(N)}{g(N)}\right| = 0$$
respectively. We also use Hardy notation: $f(N) \ll g(N)$ if $f(N) = o(g(N))$.
- Throughout, I will use ∂_i to indicate the partial derivative with respect to the i^{th} cooordinate and $\mathbf{D}$ to denote the total derivative operator: if $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$,
$$(\mathbf{DF}) = (\partial_j F_i)_{ij}$$
- $\mathbb{R}_+^K = \{\mathbf{x} \in \mathbb{R}^K : x_i \geq 0, i = 1, \ldots, K\}$.
- If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^K$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for all i, $\mathbf{x} < \mathbf{y}$ if $\mathbf{x} \leq \mathbf{y}$ and $x_i < y_i$ for at least one i, and $\mathbf{x} \lhd \mathbf{y}$ if $x_i < y_i$ for all i.
- Given a Polish space Ω, $\mathbb{D}_\Omega[0, \infty)$ denotes the space of Ω-valued càdlàg functions endowed with the Skorohod topology.
- We use $X^N \xrightarrow{\mathcal{D}} X$ to denote convergence in distribution for a sequence $\{X^N(t)\}$ of càdlàg stochastic processes,*i.e.*
$$\mathbb{E}\left[f(X^N)\right] \to \mathbb{E}\left[f(X)\right]$$
for all $f : \mathbb{D}_\Omega[0, \infty) \to \mathbb{R}$ continuous in the compact uniform topology.

2. Introduction

3. Competitive density-dependent population processes

To begin, I will introduce the object of study, intended to encompass a variety of models considered in population dynamics, population genetics, and community ecology. Throughout, I will consider

Date: May 4, 2010.

a population consisting of a finite number of discrete individuals. Each individual has a *type*, which is indexed by an integer in the set $\{1, \ldots, K\}$. All individuals of a given type are identical.

Denote the number of individuals of type $i \in \{1, \ldots, K\}$ at time t by $X_i^N(t)$ and let

$$\mathbf{X}^N(t) = \left(X_1^N(t), \ldots, X_K^N(t)\right).$$

I assume that the number of individuals of each type changes when there occurs

(1) a reproduction event in which some individual of type i produces a clutch consisting of $|\mathbf{n}| \stackrel{\text{def}}{=} n_1 + \cdots + n_K$ offspring, of which n_j are of type j, or
(2) the death of an individual of type i.

Thus $\mathbf{X}^N(t)$ is a Markov chain on $\mathbb{N}_0^K$, which may be represented as

$$\mathbf{X}^N(t) = \mathbf{X}^N(0) + \sum_{i=1}^{K} \sum_{\mathbf{n} \in \mathbb{N}_0^K} \mathbf{n} B_{i,\mathbf{n}}^N(t) - \sum_{i=1}^{K} \mathbf{e}_i D_i^N(t)$$

where $B_{i,\mathbf{n}}^N(t)$ and $D_i^N(t)$ are counting processes with intensities

$$\beta_{i,\mathbf{n}}^N \left(\frac{\mathbf{X}^N(t)}{N}\right) X_i^N(t) \quad \text{and} \quad \delta_i^N \left(\frac{\mathbf{X}^N(t)}{N}\right) X_i^N(t),$$

respectively, *i.e.*

$$\tilde{B}_{i,\mathbf{n}}^N(t) \stackrel{\text{def}}{=} B_{i,\mathbf{n}}^N(t) - \int_{0+}^{t} \beta_{i,\mathbf{n}}^N \left(\frac{\mathbf{X}^N(s)}{N}\right) X_i^N(s)\, ds \tag{1}$$

and

$$\tilde{D}_i^N(t) \stackrel{\text{def}}{=} D_i^N(t) - \int_{0+}^{t} \delta_i^N \left(\frac{\mathbf{X}^N(s)}{N}\right) X_i^N(s)\, ds$$

are martingales, with quadratic covariations

$$[\tilde{B}_{i,\mathbf{n}}^N]_t = B_{i,\mathbf{n}}^N(t) \quad \text{and} \quad [\tilde{D}_i^N]_t = D_i^N(t).$$

In this work, we focus on identifying and providing rigourous proofs of limiting processes. Elsewhere, we discuss in detail the biological implications of applications of this approach to population genetics [9].

3.1. **Assumptions.** I will assume that for all compact sets $\mathcal{K} \subseteq \mathbb{R}^K$,

$$\sum_{\mathbf{n} \in \mathbb{N}_0^K} |\mathbf{n}| \sup_{\mathbf{x} \in \mathcal{K}} \beta_{i,\mathbf{n}}^N(\mathbf{x}) < \infty \quad \text{and} \quad \sup_{\mathbf{x} \in \mathcal{K}} \sum_{\mathbf{n} \in \mathbb{N}_0^K} |\mathbf{n}|^2 \sup_{\mathbf{x} \in \mathcal{K}} \beta_{i,\mathbf{n}}^N(\mathbf{x}) \tag{2}$$

and that

$$\lim_{N \to \infty} N \sup_{\mathbf{x} \in \mathcal{K}} \left|\beta_{i,n\mathbf{e}_i}^N(\mathbf{x}) - \beta_{in}(\mathbf{x})\right| = 0 \quad \text{and} \quad \lim_{N \to \infty} N \sup_{\mathbf{x} \in \mathcal{K}} \left|\delta_i^N(\mathbf{x}) - \delta_i(\mathbf{x})\right| = 0 \tag{3}$$

uniformly on compact sets, while for all i, j and all compact $\mathcal{K} \subseteq \mathbb{R}^K$,

$$\sum_{n=1}^{\infty} \sup_{\mathbf{x} \in \mathcal{K}} n\beta_{in}(\mathbf{x}) < \infty \tag{4}$$

and

$$\sup_{\mathbf{x} \in \mathcal{K}} \sum_{\{\mathbf{n} \in \mathbb{N}_0^K : \mathbf{n} \neq |\mathbf{n}|\mathbf{e}_i\}} n_j \beta_{i,\mathbf{n}}^N(\mathbf{x}) = O\left(\tfrac{1}{N}\right) \tag{5}$$

Thus, with probability tending to 1 as $N \to \infty$, all offspring are of the same type as the parent, *i.e.* mutation is rare.

While these summability assumptions are made for technical reasons, they are eminently plausible from a biological perspective, as they simply requires that the mean and variance in the number of offspring produced in any single reproductive event are both finite. Both are readily satisfied by assuming some fixed maximal clutch size.

3.2. **Mutation.** In this section, I will introduce new notation that allows the birth-death-mutation process to more closely resemble the forms considered in the Wright-Fisher diffusion. Let

$$\hat{\beta}_i^N(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{\mathbf{n} \in \mathbb{N}_0^K} |\mathbf{n}| \, \beta_{i,\mathbf{n}}^N(\mathbf{x})$$

and

$$\mu_{ij}^N(\mathbf{x}) \stackrel{\text{def}}{=} \frac{1}{\hat{\beta}_i^N(\mathbf{x})} \sum_{\{\mathbf{n} \in \mathbb{N}_0^K : \mathbf{n} \neq |\mathbf{n}|\mathbf{e}_i\}} n_j \beta_{i,\mathbf{n}}^N(\mathbf{x}).$$

$\hat{\beta}_i^N(\mathbf{x})$ is thus the expected total reproductive output of an individual of type i per unit time in environment $\mathbf{x}$, while $\mu_{ij}^N(\mathbf{x})$ is the fraction of the expected number of offspring which are of type j. By assumption,

$$\mu_{ij}^N(\mathbf{x}) = O\left(\tfrac{1}{N}\right).$$

4. Law of Large Numbers

Let

$$\bar{\beta}_i(\mathbf{x}) = \sum_{n=1}^{\infty} n\beta_{in}(\mathbf{x})$$

and define $\mathbf{F} : \mathbb{R}^K \to \mathbb{R}^K$ by

$$F_i(\mathbf{x}) \stackrel{\text{def}}{=} \left(\bar{\beta}_i(\mathbf{x}) - \delta_i(\mathbf{x})\right) x_i.$$

Let ψ_t denote the flow of

$$\dot{\mathbf{Y}}(t) = \mathbf{F}(\mathbf{Y}(t)) \tag{6}$$

and let

$$\mathbf{Y}^N(t) = \tfrac{1}{N}\mathbf{X}^N(t).$$

We then have the following Law of Large Numbers for $\mathbf{Y}^N(t)$, which will be proven below in Section 9.1

Proposition 1. *Let $\mathbf{x} \in \mathbb{R}_+^K$, and fix $0 < r < s < 1$ and $\varepsilon > 0$ so that*

$$\mathcal{K}_{\mathbf{x},\varepsilon} \stackrel{\text{def}}{=} \overline{\{\mathbf{y} : \|\mathbf{y} - \phi_t \mathbf{x}\| < \varepsilon \quad \text{for some } t \geq 0\}} \subseteq \operatorname{int}\bigl(\mathbb{R}_+^{\mathrm{K}}\bigr).$$

Then there exists a constant $B_{\mathbf{x},\varepsilon}$ such that

$$\mathbb{P}_{\mathbf{x}} \left\{ \sup_{0 \leq \frac{1-s}{4B^2_{\mathbf{x},\varepsilon}} \ln N} \left\| \mathbf{Y}^N(t) - \psi_t \mathbf{x} \right\|^2 > N^{-r} \right\} < \frac{N^{-r}}{\varepsilon}.$$

In particular, for all $t > 0$,

$$\lim_{N \to \infty} \left\| \mathbf{Y}^N(t) - \psi_t \mathbf{x} \right\| = 0 \quad \mathbb{P}_{\mathbf{x}} - a.s.$$

Based upon this, I will say that the Markov process $\mathbf{X}^N(t)$ is *competitive* if the dynamical system (6) is competitive [5], *i.e.* $\bar{\beta}_i(\mathbf{x})$ and $\delta_i(\mathbf{x})$ are C^1 and

$$\partial_j \left(\bar{\beta}_i(\mathbf{x}) - \delta_i(\mathbf{x}) \right) \leq 0$$

for all $j \neq i$.

4.1. **Competitive Dynamical Systems.** Subsequent to Proposition 1, I will always assume $\mathbf{X}^N(t)$ is competitive, and further that the dynamical system (6):

(1) is *dissipative*: there is a compact set K that uniformly attracts each compact set of initial values,
(2) is *irreducible*: the matrix $\left(\partial_j \left(\bar{\beta}_i(\mathbf{x}) - \delta_i(\mathbf{x})\right)\right)$ is irreducible for all $\mathbf{x} \in \operatorname{int} \mathbb{R}_+^K$,
(3) has a source at the origin, and
(4) for $i \neq j$, $\partial_j \left(\bar{\beta}_i(\mathbf{x}) - \delta_i(\mathbf{x})\right) < 0$ at every equilibrium in $\mathbb{R}_+^K \setminus 0$.

Again, each of these assumptions has a biological interpretation: the first requires that the population remain finite, the second that all types interact with all other types, and the third that populations will grow when started from small initial densities. The last is required

5. Linear Birth-Death Process Approximation

6. Quasi-Neutrality and Weakly Selected Quasi-Neutrality

Henceforth, we will say that the process $\mathbf{X}^N(t)$ is *quasi-neutral* if there exist C^2 functions $\boldsymbol{\gamma} : \mathbb{R}_+^K \to \mathbb{R}_+^K$ and $R : \mathbb{R}_+ \to \mathbb{R}$ such that

$$F_i(\mathbf{x}) \stackrel{\text{def}}{=} \gamma_i(\mathbf{x}) R(\mathbf{x}). \tag{7}$$

and

(1) $\boldsymbol{\gamma}(\mathbf{0}) \rhd \mathbf{0}$,
(2) $R(\mathbf{0}) = 1$, and $\{\mathbf{x} \in \mathbb{R}_+^K : R(\mathbf{x}) \geq 0\}$ is compact, and
(3) $(DR)(\mathbf{x}) \lhd \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}_+^K$.

Under these assumptions,

$$\Omega = \{\mathbf{x} : R(\mathbf{x}) = 0\},$$

is an attracting, compact, co-dimension one C^1-submanifold of equilibria for the dynamical system (6), as can be seen by considering the Lyapunov function $V(\mathbf{x}) = R^2(\mathbf{x})$.

6.1. **Geometry of Ω.** Ω is diffeomorphic to the standard simplex

$$\Delta^{K-1} \stackrel{\text{def}}{=} \left\{ \mathbf{x} \in \mathbb{R}_+^K : \sum_{i=1}^K x_i = 1 \right\}$$

via the radial projection map

$$\boldsymbol{\rho}(\mathbf{x}) = \frac{1}{\sum_{i=1}^K x_i} \mathbf{x}.$$

This has partial derivatives

$$\partial_i \rho_k = \frac{1}{\sum_{l=1}^K x_l} (\delta_{ik} - \rho_k)$$

and

$$\partial_i \partial_j \rho_k = -\frac{1}{\left(\sum_{l=1}^K x_l\right)^2} (\delta_{ik} + \delta_{jk} - 2\rho_k)$$

and is thus differentiable when restricted to Ω, where the denominator is non-vanishing.

To see that this is invertible, consider $f_{\mathbf{x}}(t) \stackrel{\text{def}}{=} R(t\mathbf{x})$ and $F(\mathbf{x}, t) = (\mathbf{x}, f_{\mathbf{x}}(t))$. Then,

$$\det((DF)(\mathbf{x}, t)) = (\partial_t f_{\mathbf{x}})(t) = (DR)(t\mathbf{x}) \cdot \mathbf{x},$$

which, by Assumption (2) is strictly positive. Thus F^{-1} exists and is differentiable; in particular, $f_{\mathbf{x}}^{-1}(0)$ exists, is unique, and is differentiable in $\mathbf{x}$, so $\mathbf{x} \mapsto f_{\mathbf{x}}^{-1}(0)\mathbf{x} : \Delta^{K-1} \to \Omega$ is a differentiable injection.

We will henceforth write

$$n_e(\mathbf{x}) \stackrel{\text{def}}{=} f_{\mathbf{x}}^{-1}(0),$$

a naming convention whose origin will be discussed in greater detail below.

6.2. **Local Dynamics Near Ω.** Recalling (7), we see that

$$(\partial_j F_i)(\mathbf{x}) = (\partial_j \gamma_i)(\mathbf{x}) R(\mathbf{x}) x_i + \gamma_i(\mathbf{x})(\partial_j R)(\mathbf{x}) x_i + \gamma_i(\mathbf{x}) R(\mathbf{x}) \delta_{ij}.$$

In particular, writing $\mathcal{F}_i(\mathbf{x}) \stackrel{\text{def}}{=} \gamma_i(\mathbf{x}) x_i$, if $\mathbf{x}^* \in \Omega$,

$$(\mathbf{DF})(\mathbf{x}^*) = \boldsymbol{\mathcal{F}}(\mathbf{x}^*) \otimes (\mathbf{D}R)(\mathbf{x}^*),$$

Thus, $(\mathbf{DF})(\mathbf{x}^*)$ has two distinct eigenvalues, 0 and $\lambda(\mathbf{x}^*)$, where

$$\lambda(\mathbf{x}) \stackrel{\text{def}}{=} (\mathbf{D}R)(\mathbf{x}) \cdot \boldsymbol{\mathcal{F}}(\mathbf{x}). \tag{8}$$

The former has corresponding eigenspace $(\mathbf{D}R)(\mathbf{x}^*)^\perp = T_{\mathbf{x}^*}\Omega$, whilst the latter corresponds to the single eigenvector $\boldsymbol{\mathcal{F}}(\mathbf{x}^*)$.

6.3. **The Projection Map and a Time Change.** Let ψ_t denote the flow of (6), and let

$$\boldsymbol{\pi}(\mathbf{x}) \stackrel{\text{def}}{=} \lim_{t \to \infty} \psi_t \mathbf{x},$$

$$\mathcal{T}_{\mathbf{x}}(t) \stackrel{\text{def}}{=} \int_t^\infty R(\psi_s \mathbf{x})\, ds,$$

and

$$\tau(\mathbf{x}) \stackrel{\text{def}}{=} \mathcal{T}_{\mathbf{x}}(0).$$

From the definition of $\mathcal{T}_{\mathbf{x}}$, we see that

$$\mathcal{T}_{\psi_u \mathbf{x}}(t-u) = \mathcal{T}_{\mathbf{x}}(t),$$

whence $\tau(\psi_t \mathbf{x}) = \mathcal{T}_{\mathbf{x}}(t)$, a relation we shall need later.

7. Results

7.1. **Quasi-Neutral Processes.** Using these definitions, we show that $\mathbf{Y}^N$ will approach Ω-limit set with probability tending to 1 as $N \to \infty$:

Proposition 2. *Let ε and $B_{\mathbf{x},\varepsilon}$ be as in Proposition 1. Fix $r < \frac{1}{1+\frac{4B_{\mathbf{x},\varepsilon}^2}{\beta_-(1-\alpha)}}$. Then for all $r' < r < s < 1$, and N sufficiently large,*

$$\mathbb{P}_{\mathbf{x}}\left\{\left|\tau\Big(\mathbf{Y}^N\big(\tfrac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\ln N\big)\Big)\right| > N^{-r'}\right\} < \frac{N^{-r}}{\varepsilon}.$$

This establishes that the for N sufficiently large, $\mathbf{Y}^N(t)$ will come arbitrarily close to Ω. In fact, the process, having arrived at Ω, remains there. To do show this, I consider the process on a longer time-scale: let

$$\mathbf{Z}^N(t) \stackrel{\text{def}}{=} \frac{1}{N}\mathbf{X}^N\big(\tfrac{N}{2}t\big).$$

Then,

Theorem 1. *Let r' be as above, and fix $0 < \delta < r'$. Suppose $\tau(\mathbf{Z}^N(0)) < N^{-\delta}$, and let*

$$\tau_\delta = \inf\{t \geq 0 : \tau(\mathbf{Z}^N(t)) > N^{-\delta}\}.$$

Then, as $N \to \infty$, for any fixed $T > 0$,

$$\tau(\mathbf{Z}^N(T \wedge \tau_\delta)) \xrightarrow{\mathcal{D}} 0.$$

Recalling that $\mathbf{z} = \phi_{\tau(\mathbf{z})}\boldsymbol{\pi}(\mathbf{z})$, the continuous mapping theorem gives us

Corollary 1. *Under the assumptions above,*

$$\mathbf{Z}^N(t) - \boldsymbol{\pi}(\mathbf{Z}^N(t)) \xrightarrow{\mathcal{D}} 0$$

We now prepare for our main result by introducing some additional notation. Let

$$\sigma_i(\mathbf{x}) \overset{\text{def}}{=} \lim_{N\to\infty} N\left[\left(\hat{\beta}_i^N(\mathbf{x}) - \bar{\beta}_i(\mathbf{x})\right) - \left(\delta_i^N(\mathbf{x}) - \delta_i(\mathbf{x})\right)\right], \tag{9}$$

i.e. $\sigma_i(\mathbf{x})$ is the $O\left(\frac{1}{N}\right)$ component of the net reproductive rate (that the limit exists is guaranteed by Assumptions (2) – (5), let

$$\theta_{ij}(\mathbf{x}) \overset{\text{def}}{=} \lim_{N\to\infty} N\mu_{ij}^N(\mathbf{x}) \tag{10}$$

be the rescaled rate of mutation, and let

$$\check{\beta}_i(\mathbf{x}) \overset{\text{def}}{=} \sum_{n=0}^{\infty} n^2 \beta_{i,n}(\mathbf{x}).$$

With these, we have

Theorem 2. *Suppose that as $N \to \infty$, $\mathbf{Z}^N(0) \xrightarrow{\mathcal{D}} \mathbf{Z}(0) \in \Omega$, and that the Martingale problem for*

$$
\begin{aligned}
(11)\quad (Lf)(\boldsymbol{\pi}) = \sum_{i=1}^{K} \Bigg(& - \sum_{j=1}^{K} \theta_{ij}(\boldsymbol{\pi})\bar{\beta}_i(\boldsymbol{\pi})\pi_i + \sum_{j=1}^{K} \theta_{ji}(\boldsymbol{\pi})\bar{\beta}_j(\boldsymbol{\pi})\pi_i \\
& + \frac{\gamma_i(\boldsymbol{\pi})\pi_i}{\lambda(\boldsymbol{\pi})} \sum_{j=1}^{K}\sum_{k=1}^{K} \theta_{jk} \left((\partial_k R)(\boldsymbol{\pi}) - (\partial_j R)(\boldsymbol{\pi})\right) \bar{\beta}_j \pi_j \\
& + \pi_i \left(\sigma_i(\boldsymbol{\pi}) - \frac{\gamma_i(\boldsymbol{\pi})}{\lambda(\boldsymbol{\pi})} \sum_{j=1}^{K} (\partial_j R)(\boldsymbol{\pi}) \sigma_j(\boldsymbol{\pi}) \pi_j \right) \\
+ \frac{\gamma_i(\boldsymbol{\pi})\pi_i}{\lambda^2(\boldsymbol{\pi})} \Bigg(& 2 \sum_{j=1}^{K} (\partial_j R) \left((\partial_j R)(\boldsymbol{\pi})(\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi})) - (\partial_i R)(\boldsymbol{\pi})(\check{\beta}_i(\boldsymbol{\pi}) + \bar{\beta}_i(\boldsymbol{\pi})) \right) \gamma_j(\boldsymbol{\pi})\pi_j \\
& + \gamma_i(\boldsymbol{\pi}) \sum_{j=1}^{K}\sum_{k=1}^{K} (\partial_j R)(\boldsymbol{\pi})(\partial_k R)(\boldsymbol{\pi}) \partial_k \left(\frac{\gamma_j(\boldsymbol{\pi})}{\gamma_i(\boldsymbol{\pi})} \right) (\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi}))\pi_j\pi_k \\
& \sum_{j=1}^{K} (\partial_j R)^2(\boldsymbol{\pi}) \partial_j \left(\frac{\sum_{k=1}^{K} (\partial_k R)(\boldsymbol{\pi})}{(\partial_j R)(\boldsymbol{\pi})} \right) (\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi}))\pi_j \\
+ \frac{1}{\lambda(\boldsymbol{\pi})} \Bigg(& \left(\sum_{j=1}^{K} (\partial_j R)^2(\boldsymbol{\pi})(\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi}))\pi_j \right) \left(\sum_{j=1}^{K} (\partial_j R)(\boldsymbol{\pi})(\gamma_i(\boldsymbol{\pi}) - \gamma_j(\boldsymbol{\pi}))\gamma_j(\boldsymbol{\pi})\pi_j \right) \\
- \sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K} & (\partial_j R)^2(\boldsymbol{\pi}) \partial_j \left(\frac{(\partial_l R)(\boldsymbol{\pi})}{(\partial_j R)(\boldsymbol{\pi})} \right) (\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi}))\pi_j \gamma_k(\boldsymbol{\pi})\pi_k \gamma_l(\boldsymbol{\pi})\pi_l \Bigg)\Bigg)\Bigg) (\partial_i f)(\boldsymbol{\pi}) \\
+ \sum_{i=1}^{K}\sum_{j=1}^{K} \pi_i \Bigg(& \left(\check{\beta}_i(\boldsymbol{\pi}) + \bar{\beta}_i(\boldsymbol{\pi})\right) \delta_{ij} - \frac{\pi_j}{\lambda^2(\boldsymbol{\pi})} \sum_{k=1}^{K} (\partial_k R)(\boldsymbol{\pi}) \Big[\gamma_i(\boldsymbol{\pi})\gamma_k(\boldsymbol{\pi})(\partial_j R) \left(\check{\beta}_j(\boldsymbol{\pi}) + \bar{\beta}_j(\boldsymbol{\pi})\right) \\
& + \gamma_j(\boldsymbol{\pi})\gamma_k(\boldsymbol{\pi})(\partial_i R) \left(\check{\beta}_i(\boldsymbol{\pi}) + \bar{\beta}_i(\boldsymbol{\pi})\right) - \gamma_i(\boldsymbol{\pi})\gamma_j(\boldsymbol{\pi})(\partial_k R) \left(\check{\beta}_k(\boldsymbol{\pi}) + \bar{\beta}_k(\boldsymbol{\pi})\right) \Big] \pi_k \Bigg) (\partial_i \partial_j f)(\boldsymbol{\pi})
\end{aligned}
$$

is well posed. Then $\boldsymbol{\Pi}^N(t) \stackrel{\text{def}}{=} \boldsymbol{\pi}\left(\mathbf{Z}^N(t)\right)$ converges weakly to a diffusion process $\boldsymbol{\Pi}(t)$ with this generator.

In general, well-posedness for degenerate diffusions remains an open-problem, and uniqueness is not guaranteed when the coefficients fail to be Lipschitz; in section 9.5.5 I discuss sufficient conditions to ensure uniqueness.

Note that the assumption that $\mathbf{Z}^N(0) \xrightarrow{\mathcal{D}} \mathbf{Z}(0) \in \Omega$ is necessary to obtain weak convergence in the results above; Propositions 1 and 2 tell us that in general,

$$\lim_{N \to \infty} \mathbf{Z}^N(0+) = \boldsymbol{\pi}(\mathbf{Z}(0)) \neq \mathbf{Z}(0).$$

Thus, the limiting process fails to satisfy the Feller property. We may still obtain a global convergence result, by considering a variant of the original process:

Corollary 2. *Let*

$$\tilde{\mathbf{Z}}^N(t) \stackrel{\text{def}}{=} \mathbf{Z}^N(t) - \psi_{Nt}\mathbf{Z}^N(0) + \boldsymbol{\pi}(\mathbf{Z}^N(0)).$$

Then $\tilde{\mathbf{Z}}^N \xrightarrow{\mathcal{D}} \tilde{\mathbf{Z}}$, *a diffusion process with generator* (11) *as above.*

In keeping with the convention in population genetics of considering relative frequencies, rather than absolute numbers of types, we consider the process $\mathbf{P}(t) = \boldsymbol{\rho}(\Pi(t))$, where Applying Itô's formula to $\boldsymbol{\rho}(\Pi(t))$ yields the following:

Corollary 3 (Relative Frequency process)**.**

$$\begin{aligned}(12)\quad L_{RF}f(\mathbf{p}) = \sum_{i=1}^{K}\Bigg(& -\sum_{j=1}^{K}\theta_{ij}^{\rho}(\mathbf{p})\bar{\beta}_i^{\rho}(\mathbf{p})p_i + \sum_{j=1}^{K}\theta_{ji}^{\rho}(\mathbf{p})\bar{\beta}_j^{\rho}(\mathbf{p})p_i \\ & + \frac{(\gamma_i^{\rho}(\mathbf{p}) - \bar{\gamma}^{\rho}(\mathbf{p}))\,p_i}{\lambda^{\rho}(\mathbf{p})}\sum_{j=1}^{K}\sum_{k=1}^{K}\theta_{jk}^{\rho}(\mathbf{p})\left((\partial_k R)^{\rho}(\mathbf{p}) - (\partial_j R)^{\rho}(\mathbf{p})\right)\bar{\beta}_j p_j \\ & + p_i\left((\sigma_i^{\rho}(\mathbf{p}) - \bar{\sigma}^{\rho}(\mathbf{p})) - \frac{(\gamma_i^{\rho}(\mathbf{p}) - \bar{\gamma}^{\rho}(\mathbf{p}))}{\lambda^{\rho}(\mathbf{p})}\sum_{j=1}^{K}(\partial_j R)^{\rho}(\mathbf{p})\sigma_j^{\rho}(\mathbf{p})p_j\right) \\ & \cdots \\ & + \frac{1}{n_e(\mathbf{p})}\frac{(\gamma_i^{\rho}(\mathbf{p}) - \bar{\gamma}^{\rho}(\mathbf{p}))}{(\lambda^{\rho}(\mathbf{p}))^2}\sum_{j=1}^{K}\left((\partial_j R)^{\rho}(\mathbf{p})\right)^2\partial_j\left(\frac{\sum_{k=1}^{K}(\partial_k R)^{\rho}(\mathbf{p})}{(\partial_j R)^{\rho}(\mathbf{p})}\right)(\check{\beta}_j^{\rho}(\mathbf{p}) + \bar{\beta}_j^{\rho}(\mathbf{p}))p_j \\ & \cdots \\ & - \frac{(\gamma_i^{\rho}(\mathbf{p}) - \bar{\gamma}^{\rho}(\mathbf{p}))}{(\lambda^{\rho}(\mathbf{p}))^3}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{l=1}^{K}\left((\partial_j R)^{\rho}(\mathbf{p})\right)^2\partial_j\left(\frac{(\partial_l R)^{\rho}(\mathbf{p})}{(\partial_j R)^{\rho}(\mathbf{p})}\right)(\check{\beta}_j^{\rho}(\mathbf{p}) + \bar{\beta}_j^{\rho}(\mathbf{p}))p_j\gamma_k^{\rho}(\mathbf{p})p_k\gamma_l^{\rho}(\mathbf{p})p_l\Bigg)(\partial_i f)(\mathbf{p})\end{aligned}$$

Where we adopt the notation $f^{\rho}(\mathbf{p}) \stackrel{\text{def}}{=} f(\rho^{-1}(\mathbf{p}))$.

Lastly, we observe that when all types are identical to order $O\left(\frac{1}{N}\right)$, *i.e.*

$$\gamma_i(\mathbf{x}) = \gamma_j(\mathbf{x}) \quad \text{for all } 1 \leq i,j \leq K \text{ and all } \mathbf{x} \in \mathbb{R}_+^K,$$

we obtain the following process, which naturally generalises the Wright-Fisher diffusion and the diffusion approximation to Gillespie's fecundity-variance model [3, 4]:

Corollary 4 (Weak Selection)**.**

$$(13)\quad L_{WS}f(\mathbf{p}) \qquad =$$

8. Important Special Cases

8.1. **The Gause-Lotka-Volterra Model.**

8.2. **The Double Monod Model.**

8.3. **Results for $K = 2$.**

9. Proofs

9.1. Proof of Proposition 1. Define $\mathbf{F}^N : \mathbb{R}_+^K \to \mathbb{R}_+^K$ by

$$F_i^N(\mathbf{x}) = \left(\sum_{\mathbf{n} \in \mathbb{N}_0^K} n_i \beta_{i,\mathbf{n}}^N(\mathbf{x}) - \delta_i^N(\mathbf{x}) \right) x_i.$$

Thus,

$$\mathbf{X}^N(t) = \mathbf{X}^N(0) + \int_{0+}^t N\mathbf{F}^N\left(\tfrac{\mathbf{X}^N(s)}{N}\right) ds + \sum_{i=1}^K \sum_{\mathbf{n} \in \mathbb{N}_0^K} \mathbf{n} \tilde{B}_{i,\mathbf{n}}^N(t) - \sum_{i=1}^K \mathbf{e}_i \tilde{D}_i^N(t)$$

and

$$\mathbf{Y}^N(t) = \mathbf{Y}^N(0) + \int_{0+}^t \mathbf{F}^N(\mathbf{Y}^N(s))\, ds + \frac{1}{N} \sum_{i=1}^K \sum_{\mathbf{n} \in \mathbb{N}_0^K} \mathbf{n} \tilde{B}_{i,\mathbf{n}}^N(t) - \frac{1}{N} \sum_{i=1}^K \mathbf{e}_i \tilde{D}_i^N(t)$$

Let

$$A_{\mathbf{x},\varepsilon} = \sup_{\mathbf{y} \in \mathcal{K}_{\mathbf{x},\varepsilon}} \|\mathbf{y}\|.$$

Next, recall that $\mathbf{F}$ is C^2 and thus locally Lipschitz, so there exists a constant $B_{\mathbf{x},\varepsilon}$ such that

$$\|\mathbf{F}(\mathbf{y}_1) - \mathbf{F}(\mathbf{y}_2)\| < B_{\mathbf{x},\varepsilon} \|\mathbf{y}_1 - \mathbf{y}_2\|$$

for all $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{K}_{\mathbf{x},\varepsilon}$.

Now, define a stopping time

$$\tau_\varepsilon^N = \inf\{t \geq 0 : \left\|\mathbf{Y}^N(t) - \psi_t \mathbf{x}\right\| > \varepsilon\}$$

and consider the stopped process $\mathbf{Y}^N(t \wedge \tau_\varepsilon^N)$.

From **??**, we have

$$\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N,\mathbf{x})=\mathbf{Y}^N(0)-\mathbf{x}+\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}(\mathbf{Y}^N(s))-\mathbf{F}(\psi_s\mathbf{x})\,ds$$
$$+\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}^N(\mathbf{Y}^N(s))-\mathbf{F}(\mathbf{Y}^N(s))\,ds+\frac{1}{N}\left(\sum_{i=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K}\mathbf{n}\tilde{B}_{i,\mathbf{n}}^N(t\wedge\tau_\varepsilon^N)-\sum_{i=1}^{K}\mathbf{e}_i\tilde{D}_i^N(t\wedge\tau_\varepsilon^N)\right),$$

while applying the Cauchy-Schwartz inequality gives

$$\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N,\mathbf{x})\right\|^2\leq 4\Bigg(\left\|\mathbf{Y}^N(0)-\mathbf{x}\right\|^2$$
$$+\left\|\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}(\mathbf{Y}^N(s))-\mathbf{F}(\psi_s\mathbf{x})\,ds\right\|^2+\left\|\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}^N(\mathbf{Y}^N(s))-\mathbf{F}(\mathbf{Y}^N(s))\,ds\right\|^2$$
$$+\frac{1}{N^2}\left\|\sum_{i=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K}\mathbf{n}\tilde{B}_{i,\mathbf{n}}^N(t\wedge\tau_\varepsilon^N)-\sum_{i=1}^{K}\mathbf{e}_i\tilde{D}_i^N(t\wedge\tau_\varepsilon^N)\right\|^2\Bigg).$$

Applying Jensen's inequality gives

$$\sup_{t\leq T}\left\|\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}(\mathbf{Y}^N(s))-\mathbf{F}(\psi_s\mathbf{x})\,ds\right\|^2\leq B_{\mathbf{x},\varepsilon}^2T\int_{0+}^{T\wedge\tau_\varepsilon^N}\left\|\mathbf{Y}^N(s))-\psi_s\mathbf{x}\right\|\,ds$$

and

$$\sup_{t\leq T}\left\|\int_{0+}^{t\wedge\tau_\varepsilon^N}\mathbf{F}^N(\mathbf{Y}^N(s))-\mathbf{F}(\mathbf{Y}^N(s))\,ds\right\|^2\leq T\int_{0+}^{T\wedge\tau_\varepsilon^N}\left\|\mathbf{F}^N(\mathbf{Y}^N(s))-\mathbf{F}(\mathbf{Y}^N(s))\right\|^2\,ds,$$

while

$$F_i^N(\mathbf{y})-F_i(\mathbf{y})=\sum_{n=1}^{\infty}n\left(\beta_{i,n\mathbf{e}_i}^N(\mathbf{y})-\beta_{in}(\mathbf{y})\right)y_i(s)$$
$$+\sum_{j=1}^{K}\sum_{\{\mathbf{n}\in\mathbb{N}_0^K:\mathbf{n}\neq|\mathbf{n}|\mathbf{e}_i\}}n_i\beta_{j,\mathbf{n}}^N(\mathbf{y})y_j(s)-\left(\delta_i^N(\mathbf{y})-\delta_i(\mathbf{y})\right)y_i(s),$$

so by assumption, we have $\left\|\mathbf{F}^N(\mathbf{y})-\mathbf{F}(\mathbf{y})\right\|=O\left(\frac{1}{N}\right)$ uniformly on $\mathcal{K}_{\mathbf{x},\varepsilon}$.

Lastly, Doob's inequality gives

$$\mathbb{E}\left[\sup_{t\leq T}\left|\sum_{j=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K}n_i\tilde{B}_{j,\mathbf{n}}^N(t\wedge\tau_\varepsilon^N)-\tilde{D}_i^N(t\wedge\tau_\varepsilon^N)\right|^2\right]\leq 4\mathbb{E}\left[\left[\sum_{j=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K}n_i\tilde{B}_{j,\mathbf{n}}^N-\tilde{D}_i^N\right](T\wedge\tau_\varepsilon^N)\right]$$
$$\leq N\int_{0+}^{T\wedge\tau_\varepsilon^N}\mathbb{E}\left[\left|\sum_{j=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K}n_i\beta_{j,\mathbf{n}}^N(\mathbf{Y}^N(s))Y_i^N(s)+\delta_i^N(\mathbf{Y}^N(s))Y_i^N(s)\right|\right]\,ds,$$

and again, assumptions made on the intensities insure that $\left|\sum_{j=1}^{K}\sum_{\mathbf{n}\in\mathbb{N}_0^K} n_i\beta_{j,\mathbf{n}}^N(\mathbf{y})y_i + \delta_i^N(\mathbf{y})y_i\right|$ is uniformly bounded on $\mathcal{K}_{\mathbf{x},\varepsilon}$.

Combining these, we have

$$\mathbb{E}\left[\sup_{t\le T}\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N)\right\|^2\right]$$
$$\le 4\mathbb{E}\left[\left\|\mathbf{Y}^N(0)-\mathbf{Y}(0)\right\|^2\right]+\frac{C_{\mathbf{x},\varepsilon}T}{N}+4B_{\mathbf{x},\varepsilon}^2 T\int_0^{T\wedge\tau_\varepsilon^N}\mathbb{E}\left[\left\|\mathbf{Y}^N(s)-\mathbf{Y}(s)\right\|^2\right]ds$$

for a constant $C_{\mathbf{x},\varepsilon}$. Applying Gronwall's inequality, we have

$$\mathbb{E}\left[\sup_{t\le T}\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N)\right\|^2\right]\le\left(4\mathbb{E}\left[\left\|\mathbf{Y}^N(0)-\mathbf{Y}(0)\right\|^2\right]+\frac{C_{\mathbf{x},\varepsilon}T}{N}\right)e^{4B_{\mathbf{x},\varepsilon}^2T}, \tag{14}$$

from which, taking $T=\frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\ln N$, we obtain

$$\sup_{t\le\frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\ln N}\mathbb{E}_{\mathbf{x}}\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N)\right\|^2\le\frac{C_{\mathbf{x},\varepsilon}(1-s)}{4B_{\mathbf{x},\varepsilon}}(\ln N)\,N^{-s}<N^{-r}$$

for N sufficiently large.

Lastly, we observe

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\left\{t\wedge\tau_\varepsilon^N<t\right\}&=\mathbb{P}_{\mathbf{x}}\left\{\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N)\right\|^2\ge\varepsilon\right\}\\&\le\frac{\mathbb{E}_{\mathbf{x}}\left\|\mathbf{Y}^N(t\wedge\tau_\varepsilon^N)-\mathbf{Y}(t\wedge\tau_\varepsilon^N)\right\|^2}{\varepsilon}\\&<\frac{N^{-r}}{\varepsilon},\end{aligned}$$

from which our result follows.

9.2. **The Projection Map and its Derivatives.** We begin by considering a new dynamical system, with trajectories identical to (6), but traversed backwards in time, away from Ω.

Lemma 1. *Let ϕ_t be the flow of*

$$\dot{\boldsymbol{\mathcal{Y}}}=-\boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{Y}}) \tag{15}$$

then

$$\psi_t\mathbf{x}=\phi_{\mathcal{T}_{\mathbf{x}}(t)}\boldsymbol{\pi}(\mathbf{x}). \tag{16}$$

Proof. Recall that

$$\mathcal{T}_{\mathbf{x}}(t)=\int_t^\infty R(\psi_s\mathbf{x})\,ds.$$

First, assume that $\mathbf{x}$ is a rest point for (6); then, for all t, $\psi_t\mathbf{x} = \mathbf{x}$ and $R(\psi_t\mathbf{x}) = 0$. Thus, $\mathcal{T}_{\mathbf{x}}(t) \equiv 0$ and

$$\psi_t\mathbf{x} = \mathbf{x} = \boldsymbol{\pi}(\mathbf{x}) = \phi_{\mathcal{T}_{\mathbf{x}}(t)}\boldsymbol{\pi}(\mathbf{x})$$

as desired.

If $\mathbf{x}$ is not a rest point, then $R(\psi_t\mathbf{x})$ is strictly positive or strictly negative for all $t \in \mathbb{R}$, and thus the inverse function $\mathcal{T}_{\mathbf{x}}^{-1}$ is everywhere defined and has derivative

$$\frac{d}{dt}\mathcal{T}_{\mathbf{x}}^{-1}(t) = -\frac{1}{R(\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x})}.$$

Thus

$$\frac{d}{dt}\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x} = \frac{1}{R(\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x})}\mathbf{F}(\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x}) = \boldsymbol{\mathcal{F}}(\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x}),$$

while

$$\lim_{t\to 0}\psi_{\mathcal{T}_{\mathbf{x}}^{-1}(t)}\mathbf{x} = \lim_{t\to\infty}\psi_t\mathbf{x} = \boldsymbol{\pi}(\mathbf{x}).$$

The result then follows from uniqueness of solutions. $\square$

Henceforth, I will use ϕ_t in favour of ψ_t when considering the trajectories of (6). In particular, taking $t = 0$ in (16) gives the following essential identity:

$$\mathbf{x} = \phi_{\tau(\mathbf{x})}\boldsymbol{\pi}(\mathbf{x}) \quad \text{and} \quad \boldsymbol{\pi}(\mathbf{x}) = \phi_{-\tau(\mathbf{x})}\mathbf{x}, \tag{17}$$

As a first application, this may be used to find the derivatives of $\boldsymbol{\pi}(\mathbf{x})$ and $\tau(\mathbf{x})$. Differentiating (17) yields

$$(\mathbf{D}\boldsymbol{\pi})(\mathbf{x}) = e^{-\int_{0+}^{-\tau(\mathbf{x})}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds} + \boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x}))(\mathbf{D}\tau)(\mathbf{x})$$

where we have exploited the identities [6]

$$\frac{d}{dt}\phi_t\mathbf{x} = -\boldsymbol{\mathcal{F}}(\phi_t\mathbf{x}) \quad \text{and} \quad (\mathbf{D}\phi_t)(\mathbf{x}) = e^{-\int_{0+}^{t}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds}.$$

Differentiating $R(\boldsymbol{\pi}(\mathbf{x})) = 0$, we have

$$\begin{aligned}\mathbf{0} &= (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}))(\mathbf{D}\boldsymbol{\pi})(\mathbf{x}) \\ &= (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}))\left(e^{-\int_{0+}^{-\tau(\mathbf{x})}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds} + \boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x}))(\mathbf{D}\tau)(\mathbf{x})\right)\end{aligned}$$

whence

$$(\mathbf{D}\tau)(\mathbf{x}) = -\frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))}(\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}))e^{-\int_{0+}^{-\tau(\mathbf{x})}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds}, \tag{18}$$

Substituting this into the expression for $(\mathbf{D}\boldsymbol{\pi})(\mathbf{x})$, above, we have

$$\begin{aligned}(\mathbf{D}\boldsymbol{\pi})(\mathbf{x}) &= \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))}\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x}))(\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}))\right)e^{-\int_{0+}^{-\tau(\mathbf{x})}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds} \\ &= \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))}(\mathbf{D}\mathbf{F})(\boldsymbol{\pi}(\mathbf{x}))\right)e^{-\int_{0+}^{-\tau(\mathbf{x})}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds}.\end{aligned} \tag{19}$$

To find the second derivatives, recall that the matrix exponential has Fréchet derivative [1]

$$\lim_{h\downarrow 0} \frac{e^{A+hB} - e^A}{h} = \int_{0+}^1 e^{uA} B e^{(1-u)A}\, du,$$

so that

$$\partial_i \left(e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \right) =$$

$$\int_{0+}^1 e^{u \int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \left(-(\mathbf{D}\boldsymbol{\mathcal{F}})(\boldsymbol{\pi}(\mathbf{x}))(\partial_i \tau)(\mathbf{x}) + \int_{0+}^{-\tau(\mathbf{x})} \partial_i \left[(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x}) \right] ds \right) e^{(1-u) \int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds}\, du.$$

Some matrix manipulations give

$$\begin{aligned}
\partial_i \left(\lambda(\boldsymbol{\pi}(\mathbf{x})) \right) &= \left(\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x}))^t (\mathbf{D}^2 R)(\boldsymbol{\pi}(\mathbf{x})) + (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}))(\mathbf{D}\boldsymbol{\mathcal{F}})(\boldsymbol{\pi}(\mathbf{x})) \right) (\partial_i \boldsymbol{\pi})(\mathbf{x}) \\
&= \sum_{k=1}^K \sum_{l=1}^K \left((\partial_{kl} R)(\boldsymbol{\pi}(\mathbf{x})) \mathcal{F}_k(\boldsymbol{\pi}(\mathbf{x})) + (\partial_k R)(\boldsymbol{\pi}(\mathbf{x}))(\partial_l \mathcal{F}_k)(\boldsymbol{\pi}(\mathbf{x})) \right) (\partial_i \pi_l)(\mathbf{x})
\end{aligned}$$

$$\begin{aligned}
(\partial_i \mathbf{D}\tau)(\mathbf{x}) = &\left(-\frac{\partial_i \left(\lambda(\boldsymbol{\pi}(\mathbf{x})) \right)}{\lambda^2(\boldsymbol{\pi}(\mathbf{x}))} (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x})) + \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))} (\mathbf{D}^2 R)(\boldsymbol{\pi}(\mathbf{x}))(\partial_i \boldsymbol{\pi})(\mathbf{x}) \right) e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \\
&+ \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))} (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x})) \partial_i \left(e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \right)
\end{aligned}$$

$$\begin{aligned}
(\partial_i \mathbf{D}\boldsymbol{\pi})(\mathbf{x}) &= \partial_i \left(e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \right) - (\mathbf{D}\boldsymbol{\mathcal{F}})(\boldsymbol{\pi}(\mathbf{x}))(\partial_i \boldsymbol{\pi})(\mathbf{x}) \otimes (\mathbf{D}\tau)(\mathbf{x}) - \boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x})) \otimes (\partial_i \mathbf{D}\tau)(\mathbf{x}) \\
&= \Bigg(\frac{\partial_i \left(\lambda(\boldsymbol{\pi}(\mathbf{x})) \right)}{\lambda^2(\boldsymbol{\pi}(\mathbf{x}))} (\mathbf{DF})(\boldsymbol{\pi}(\mathbf{x})) \\
&\qquad + \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))} \left((\mathbf{D}\boldsymbol{\mathcal{F}})(\boldsymbol{\pi}(\mathbf{x}))(\partial_i \boldsymbol{\pi})(\mathbf{x}) \otimes (\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{x}) + \boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{x})) \otimes (\mathbf{D}^2 R)(\boldsymbol{\pi}(\mathbf{x}))(\partial_i \boldsymbol{\pi})(\mathbf{x}) \right) \Bigg) \\
&\qquad \times e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} + \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{x}))} (\mathbf{DF})(\boldsymbol{\pi}(\mathbf{x})) \right) \partial_i \left(e^{\int_{0+}^{-\tau(\mathbf{x})} (\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s \mathbf{x})\, ds} \right)
\end{aligned}$$

The various partial derivatives can then be obtained via the identities $(\partial_i \tau)(\mathbf{x}) = (\mathbf{D}\tau)(\mathbf{x})\mathbf{e}_i$, $(\partial_i \pi_k) = \mathbf{e}_j \cdot (\mathbf{D}\boldsymbol{\pi})(\mathbf{x})\mathbf{e}_k$, *et cetera.*

To end this section, we develop several other identities that will be useful: first, note that

$$\frac{d}{dt} R(\phi_t \mathbf{x}) = -(\mathbf{D}R)(\phi_t \mathbf{x}) \cdot \boldsymbol{\mathcal{F}}(\phi_t \mathbf{x}) = -\lambda(\phi_t \mathbf{x}),$$

so that

$$R(\psi_t \mathbf{x}) = R(\phi_{\mathcal{T}_{\mathbf{x}}(t)} \boldsymbol{\pi}(\mathbf{x})) - R(\boldsymbol{\pi}(\mathbf{x})) = -\int_{0+}^{\mathcal{T}_{\mathbf{x}}(t)} \lambda(\phi_s \boldsymbol{\pi}(\mathbf{x}))\, ds, \tag{20}$$

and, in particular,

$$R(\mathbf{x}) = -\int_{0+}^{\tau(\mathbf{x})} \lambda(\phi_s \boldsymbol{\pi}(\mathbf{x}))\, ds. \tag{21}$$

Secondly,

$$\frac{d}{dt}\boldsymbol{\mathcal{F}}(\phi_t\mathbf{x}) = -(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_t\mathbf{x})\boldsymbol{\mathcal{F}}(\phi_t\mathbf{x}),$$

whence

$$\boldsymbol{\mathcal{F}}(\phi_t\mathbf{x}) = e^{-\int_{0+}^{t}(\mathbf{D}\boldsymbol{\mathcal{F}})(\phi_s\mathbf{x})\,ds}\boldsymbol{\mathcal{F}}(\mathbf{x}). \tag{22}$$

9.3. **Proof of Propostion 2.** By the defintiion of $\mathcal{T}_{\mathbf{x}}(t)$ and (20), we have

$$\partial_t\mathcal{T}_{\mathbf{x}}(t) = -R(\psi_t\mathbf{x}) = \int_{0+}^{\mathcal{T}_{\mathbf{x}}(t)}\lambda(\phi_s\boldsymbol{\pi}(\mathbf{x}))\,ds \leq \lambda_{\mathbf{x},\varepsilon}\mathcal{T}_{\mathbf{x}}(t).$$

where

$$\lambda_{\mathbf{x},\varepsilon} \stackrel{\text{def}}{=} \sup_{\mathbf{y}\in\mathcal{K}_{\mathbf{x},\varepsilon}}\lambda(\mathbf{y}) < 0.$$

Applying Grönwall's Inequality gives

$$\mathcal{T}_{\mathbf{x}}(t) \leq e^{\lambda_{\mathbf{x},\varepsilon}t}\mathcal{T}_{\mathbf{x}}(0),$$

or, equivalently, $\tau(\psi_t\mathbf{x}) \leq e^{\lambda_{\mathbf{x},\varepsilon}t}\tau(\mathbf{x})$. Taking $t = \frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\ln N$, we have

$$\tau(\psi_t\mathbf{x}) \leq N^{\lambda_{\mathbf{x},\varepsilon}\frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}}\tau(\mathbf{x}).$$

Lastly $\tau(\mathbf{x})$ is C^2 and thus locally Lipschitz, so

$$\left|\tau(\mathbf{Y}^N(t)) - \tau(\psi_t\mathbf{x})\right| \leq L\left\|\mathbf{Y}^N(t) - \psi_t\mathbf{x}\right\|.$$

Thus, provided $r < \frac{|\lambda_{\mathbf{x},\varepsilon}|}{|\lambda_{\mathbf{x},\varepsilon}|+12B_{\mathbf{x},\varepsilon}^2}$ and $r' < r < s < 1$, then $-r < \lambda_{\mathbf{x},\varepsilon}\frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}$, and, if N is sufficiently large,

$$\left|\tau\left(\mathbf{Y}^N(\tfrac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\ln N)\right)\right| > N^{-r'}$$

only if $\left\|\mathbf{Y}^N(t) - \psi_t\mathbf{x}\right\| > N^{-r}$. The result then follows from Proposition 1.

9.4. **Proof of Theorem 1.** Our strategy will be to show that $\tau^2(\mathbf{Z}^N(t\wedge\tau_\delta))$ converges weakly to 0. Then

$$\mathbb{P}\left\{t\wedge\tau_\delta < t\right\} = \mathbb{P}\left\{\tau^2(\mathbf{Z}^N(t\wedge\tau_\delta)) \geq \delta\right\} \leq \frac{\mathbb{E}\left[\tau^2(\mathbf{Z}^N(t\wedge\tau_\delta))\right]}{\delta} \to 0$$

as $N\to\infty$. This allows us to conclude that $\tau(\mathbf{Z}^N(t)) \xrightarrow{\mathcal{D}} 0$. We begin with a few preliminaries that are essential to our argument.

First, we recall that we have assumed that $\lambda(\mathbf{x}^*) < 0$ for all points $\mathbf{x}^*\in\Omega$. Since Ω is compact, we can fix $\delta > 0$ such that $\lambda(\mathbf{x}) < 0$ for all $\mathbf{x}\in\Omega_\delta$,

$$\Omega_\delta \stackrel{\text{def}}{=} \bigcup_{t\in[-\delta,\delta]}\phi_t(\Omega).$$

By virtue of the joint continuity of $\phi_t\mathbf{x}$ in t and $\mathbf{x}$ and the compactness of Ω, Ω_δ is a compact set.

Let

$$\tau_\delta \stackrel{\text{def}}{=} \inf\{t\geq 0 : \left|\tau(\mathbf{Z}^N(s))\right| \geq \delta\}$$

Then, (17) gives

$$\mathbf{Z}^N(t\wedge\tau_\delta)=\phi_{\tau(\mathbf{Z}^N(t\wedge\tau_\delta))}\boldsymbol{\pi}(\mathbf{Z}^N(t\wedge\tau_\delta))\in\Omega_\delta$$

i.e. $\mathbf{Z}^N(t\wedge\tau_\delta)$ is compactly contained for all $t>0$ and all N.

Next, as before, we have the following integral equation for $\mathbf{Z}^N(t)$:

$$\mathbf{Z}^N(t)=\mathbf{Z}^N(0)+\int_{0+}^t N\mathbf{F}^N(\mathbf{Z}^N(s-))\,ds+\mathbf{M}^N(t), \tag{23}$$

where

$$\mathbf{M}^N(t)\stackrel{\text{def}}{=}\frac{1}{N}\left(\sum_{i=1}^K\sum_{\mathbf{n}\in\mathbb{N}_0^K}\mathbf{n}\tilde{B}_{i,\mathbf{n}}^N(Nt)-\sum_{i=1}^K\mathbf{e}_i\tilde{D}_i^N(Nt)\right)$$

is a square integrable martingale with component-wise quadratic variations

$$\left[M_i^N\right]_t=\frac{1}{N^2}\left(\sum_{\mathbf{n}\in\mathbb{N}_0^K}n_i^2B_{i,\mathbf{n}}^N(Nt)+D_i^N(Nt)\right),$$

while for $i\neq j$, $[M_i^N,M_j^N]_t=0.$ and corresponding Meyer process

$$\left\langle M_i^N\right\rangle_t=\int_{0+}^t\sum_{\mathbf{n}\in\mathbb{N}_0^K}n_i\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s-))Z_i^N(s-)+\sum_{i=1}^K\delta_i^N(\mathbf{Z}^N(s-))Z_i^N(s-)\,ds. \tag{24}$$

In particular, we recall that by our assumption, $\sup_{N\in\mathbb{N}_0}\sum_{\mathbf{n}\in\mathbb{N}_0^K}\mathbf{n}\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s))$ and $\sup_{N\in\mathbb{N}_0}\delta_i^N(\mathbf{Z}^N(s))$ are uniformly bounded on Ω_δ. Thus, for $t\le\tau_\delta$, there exists a constant C_ε such that

$$\left\langle M_i^N\right\rangle(t)<C_\varepsilon t. \tag{25}$$

We now turn to the core of our argument. Since $\mathbf{Z}^N(t)$ is a quadratic pure jump process for all N, Itô's Formula takes the following simple form when applied to τ [10]

$$\begin{aligned}\tau(\mathbf{Z}^N(t))=\tau(\mathbf{Z}^N(0))&+\sum_{i=1}^K\int_{0+}^t(\partial_i\tau)(\mathbf{Z}^N(s-))\,dZ_i^N(s)\\&+\sum_{0<s\le t}\left\{\tau(\mathbf{Z}^N(s))-\tau(\mathbf{Z}^N(s-))-\sum_{i=1}^K(\partial_i\tau)(\mathbf{Z}^N(s-))\Delta Z_i^N(s)\right\}\end{aligned} \tag{26}$$

Formally expanding the he right hand side of (26) in powers of N, and recalling that $\mathbf{F}^N(\mathbf{x})-\mathbf{F}(x)=O\left(\frac{1}{N}\right)$ uniformly on on compacts, we see that the highest order term is

$$\begin{aligned}&N\int_{0+}^t(\mathbf{D}\tau)(\mathbf{Z}^N(s-))\cdot\mathbf{F}(\mathbf{Z}^N(s))\,ds\\&\quad=-N\int_{0+}^t\frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s)))}(\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{Z}^N(s)))e^{\int_{0+}^{-\tau(\mathbf{Z}^N(s))}(\mathbf{D}\boldsymbol{\mathfrak{F}})(\phi_u\mathbf{Z}^N(s))\,du}\boldsymbol{\mathfrak{F}}(\mathbf{Z}^N(s))R(\mathbf{Z}^N(s))\,ds\end{aligned}$$

which, using (22) simplifies to

$$-N\int_{0+}^{t}\frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s)))}(\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{Z}^N(s)))\mathfrak{F}(\boldsymbol{\pi}(\mathbf{Z}^N(s)))R(\mathbf{Z}^N(s))\,ds = -N\int_{0+}^{t}R(\mathbf{Z}^N(s))\,ds$$

Lastly, from (21), this is

$$N\int_{0+}^{t}\int_{0+}^{\tau(\mathbf{Z}^N(s))}\lambda(\phi_u\boldsymbol{\pi}(\mathbf{Z}^N(s)))\,du\,ds,$$

For $t<\tau_\delta$ and N sufficiently large, there exists $\lambda^+>0$ such that

$$\lambda(\phi_u\boldsymbol{\pi}(\mathbf{Z}^N(s)))<-\lambda^+<0,$$

so

$$N\int_{0+}^{t}\lambda^+\tau(\mathbf{Z}^N(s))\,ds < -N\int_{0+}^{t}\int_{0+}^{\tau(\mathbf{Z}^N(s))}\lambda(\phi_u\boldsymbol{\pi}(\mathbf{Z}^N(s)))\,du\,ds.$$

We proceed by imitating the proof of linearised stability [6], where $-N\lambda^+$ acts in the role of an upper bound to the eigenvalues of the Jacobian, To this end, consider

$$e^{2N\lambda^+t}\tau(\mathbf{Z}^N(t))^2.$$

Again using Itô's formula to this function, we have, after some simplification,

$$\begin{aligned}
(27)\quad \tau(\mathbf{Z}^N(t))^2 &= e^{-2N\lambda^+t}\tau(\mathbf{Z}^N(0))^2\\
&+\int_{0+}^{t}2Ne^{2N\lambda^+(s-t)}\left(\lambda^+\tau(\mathbf{Z}^N(s-))+\int_{0+}^{t}\int_{0+}^{\tau(\mathbf{Z}^N(s))}\lambda(\phi_u\boldsymbol{\pi}(\mathbf{Z}^N(s)))\,du\,ds\right)\tau(\mathbf{Z}^N(s-))\,ds\\
&+\int_{0+}^{t}2Ne^{2N\lambda^+(s-t)}\tau(\mathbf{Z}^N(s-))(\mathbf{D}\tau)(\mathbf{Z}^N(s))\cdot\left(\mathbf{F}^N(\mathbf{Z}^N(s))-\mathbf{F}(\mathbf{Z}^N(s))\right)\,ds\\
&+\int_{0+}^{t}2e^{2N\lambda^+(s-t)}\tau(\mathbf{Z}^N(s-))(\partial_i\tau)(\mathbf{Z}^N(s-))\,d\mathbf{M}^N(s)+\epsilon^N(t),
\end{aligned}$$

where

$$\begin{aligned}
(28)\quad \epsilon^N(t) = \sum_{0<s\le t}e^{2N\lambda^+(s-t)}\Big\{&\tau(\mathbf{Z}^N(s))^2-\tau(\mathbf{Z}^N(s-))^2\\
&-\sum_{i=1}^{K}2\tau(\mathbf{Z}^N(s-))(\partial_i\tau)(\mathbf{Z}^N(s-))\Delta Z_i^N(s)\Big\}.
\end{aligned}$$

By our choice of λ^+ the second term on the right is always non-positive, so

$$\begin{aligned}
\sup_{t\le T}\tau(\mathbf{Z}^N(t\wedge\tau_\delta))^2 &\le e^{-2N\lambda^+T\wedge\tau_\delta}\tau(\mathbf{Z}^N(0))^2\\
&+\int_{0+}^{T\wedge\tau_\delta}2Ne^{2N\lambda^+(s-t)}\tau(\mathbf{Z}^N(s-))(\mathbf{D}\tau)(\mathbf{Z}^N(s))\cdot\left(\mathbf{F}^N(\mathbf{Z}^N(s))-\mathbf{F}(\mathbf{Z}^N(s))\right)\,ds\\
&+\int_{0+}^{T\wedge\tau_\delta}2e^{2N\lambda^+(s-T)}\tau(\mathbf{Z}^N(s-))(\partial_i\tau)(\mathbf{Z}^N(s-))\,d\mathbf{M}^N(s)+\epsilon^N(t).
\end{aligned}$$

By assumption the first converges in distribution to 0. Theorem 4.1 in [2] tells us to show weak convergence to 0, it suffices to show that the remaining terms converge to 0 in probability.

We have assumed that $\mathbf{F}^N(\mathbf{x}) - \mathbf{F}(x) = O\big(\frac{1}{N}\big)$ uniformly on on compacts, so the third term is bounded above by a constant multiple of

$$\int_{0+}^{t} e^{2N\lambda^+(s-t)}\, ds = \frac{1}{2N\lambda^+}\Big(1 - e^{-2N\lambda^+ t}\Big).$$

The fourth term is a square-integrable martingale with Meyer process

$$N\int_{0+}^{t} \Big(2e^{2N\lambda^+(s-t)}\tau(\mathbf{Z}^N(s-))(\partial_i\tau)(\mathbf{Z}^N(s-))\Big)^2\, d\left\langle \mathbf{M}^N\right\rangle(s).$$

We recall Lenglart's inequality [8]: for $0 < p \leq 2$, there exists a constant C_p such that

$$\mathbb{E}\left[\sup_{t\leq T\wedge\tau_\delta} \left|M^N(s)\right|^p\right] \leq C_p\mathbb{E}\left[\left\langle M^N\right\rangle^{\frac{p}{2}}(T\wedge\tau_\delta)\right].$$

In particular, when $p = 1$, the above, combined with Jensen's inequality gives

$$\mathbb{E}\left[\sup_{t\leq T\wedge\tau_\delta} \left|M^N(s)\right|\right] \leq C_p\mathbb{E}\left[\left\langle M^N\right\rangle(T\wedge\tau_\delta)\right]^{\frac{1}{2}}.$$

Applying this, and remembering that τ, $\partial_i\tau$ are differentiable, and thus uniformly bounded on compacts, we see that the third integral is bounded by a constant multiple of

$$\mathbb{E}\left[N\int_{0+}^{T\wedge\tau_\delta} e^{2N\lambda^+(s-T\wedge\tau_\delta)} d\left\langle \mathbf{M}^N\right\rangle(s)\right]^{\frac{1}{2}},$$

which, by (25), is bounded above by a constant multiple of

$$\left(\int_{0+}^{T} e^{2N\lambda^+(s-T\wedge\tau_\delta)}\, ds\right)^{\frac{1}{2}}.$$

Lastly, Taylor's theorem gives

$$\tau(\mathbf{Z}^N(s))^2 - \tau(\mathbf{Z}^N(s-))^2 - \sum_{i=1}^{K} 2\tau(\mathbf{Z}^N(s-))(\partial_i\tau)(\mathbf{Z}^N(s-))\Delta Z_i^N(s)$$
$$= \sum_{i=1}^{K}\sum_{j=1}^{K}(\partial_i\tau\partial_j\tau + \tau\partial_i\partial_j\tau)(\boldsymbol{\zeta})\Delta Z_i^N(s)\Delta Z_j^N(s)$$

for some $\boldsymbol{\zeta}$ such that $\left\|\mathbf{Z}^N(s-) - \boldsymbol{\zeta}\right\| \leq \left\|\Delta\mathbf{Z}^N(s)\right\| \leq \frac{1}{N}$, so, as before, there is a constant C such that this is bounded above by

$$C\sum_{i=1}^{K}\sum_{j=1}^{K}\left|\Delta Z_i^N(s)\right|\left|\Delta Z_j^N(s)\right| \leq CK\sum_{i=1}^{K}\left|\Delta Z_i^N(s)\right|^2.$$

Thus, since $\mathbf{Z}^N$ is a quadratic pure-jump process,

$$\left|\epsilon^N(t)\right| \leq CK\sum_{0<s\leq t}\sum_{i=1}^{K} e^{2N\lambda^+(s-t)}\left|\Delta Z_i^N(s)\right|^2 = CK\sum_{i=1}^{K}\int_{0+}^{t} e^{2N\lambda^+(s-t)}\, d\left[M_i^N\right]_s.$$

and, as above,

$$\mathbb{E}\left[\sup_{t\le T\wedge\tau_\delta}\left|\epsilon^N(t)\right|\right]\le C'\left(\int_{0+}^t e^{2N\lambda^+(s-t)}\,ds\right)^{\frac{1}{2}},$$

for some constant C'.

Combining all the above, we see that

$$\mathbb{E}\left[\sup_{t\le T\wedge\tau_\delta}\tau^2(\mathbf{Z}^N(t))\right]=O\Big(\tfrac{1}{\sqrt{N}}\Big).$$

9.5. **Proof of Theorem 2.** Theorem 2 is a consequence of Theorem 5.4 in [7]; we begin Itô's formula to $\pi_i(\mathbf{Z}^N(t))$, from which we obtain an SDE for $\mathbf{\Pi}^N(t)$.

$$\begin{aligned}\pi_i(\mathbf{Z}^N(t)) &= \pi_i(\mathbf{Z}^N(0))+\sum_{i=1}^K\int_{0+}^t(\partial_i\pi_i)(\mathbf{Z}^N(s-))\,dZ_i^N(s-)\\ &\quad+\frac{1}{2}\sum_{1\le i,j\le K}\int_{0+}^t(\partial_i\partial_j\pi_i)(\mathbf{Z}^N(s-))d[Z_i^N,Z_j^N]_s+\eta^N(t)\\ &=\pi_i(\mathbf{Z}^N(0))+N\int_{0+}^t(\mathbf{D}\pi_i)(\mathbf{Z}^N(s-))\cdot F^N(\mathbf{Z}^N(s-))\,ds+\sum_{i=1}^K\int_{0+}^t(\partial_i\pi_i)(\mathbf{Z}^N(s-))\,dM_i^N(s)\\ &\quad+\sum_{i=1}^K\int_{0+}^t(\partial_i^2\pi_i)(\mathbf{Z}^N(s-))d[M_i^N]_s+\varepsilon^N(t)\end{aligned}\tag{29}$$

where

$$\begin{aligned}\eta^N(t)=\sum_{0<s\le t}\Big\{&\pi_i(\mathbf{Z}^N(s))-\pi_i(\mathbf{Z}^N(s-))-\sum_{i=1}^K\partial_i\pi_i(\mathbf{Z}^N(s-))\Delta Z_i^N(s)\\ &-\frac{1}{2}\sum_{1\le i,j\le K}(\partial_i\partial_j\pi_i)(\mathbf{Z}^N(s-))\Delta Z_i^N(s)\Delta Z_j^N(s)\Big\}.\end{aligned}\tag{30}$$

We consider each of the components individually:

9.5.1. *Finite variation terms.* The first three lines in (11) arise from considering those terms in (29) arising from the finite variation component,

$$N\int_{0+}^t(\mathbf{D}\pi_i)(\mathbf{Z}^N(s-))\cdot F^N(\mathbf{Z}^N(s-))\,ds.$$

By definition, we have

$$NF_i^N(\mathbf{Z}^N(s-)) = N\left[\sum_{\mathbf{n}\in\mathbb{N}_0^K} n_i\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s-))Z_i^N(s-) - \delta_i^N(\mathbf{x})Z_i^N(s-)\right].$$

Gathering terms appropriately, this equals

$$-\sum_{j=1}^{K} N\mu_{ij}^N(\mathbf{Z}^N(s-))\hat{\beta}_i^N(\mathbf{Z}^N(s-))Z_i^N(s-) + \sum_{j=1}^{K} N\mu_{ji}^N(\mathbf{Z}^N(s-))\hat{\beta}_j^N(\mathbf{Z}^N(s-))Z_i^N(s-) \tag{31}$$

$$+N\left[\left(\hat{\beta}_i^N(\mathbf{Z}^N(s-)) - \bar{\beta}_i(\mathbf{Z}^N(s-))\right) - \left(\delta_i^N(\mathbf{Z}^N(s-)) - \delta_i(\mathbf{Z}^N(s-))\right)\right] \tag{32}$$

$$+NF_i(\mathbf{Z}^N(s-)). \tag{33}$$

Now, recall (19), we have

$$(\mathbf{D}\boldsymbol{\pi})(\mathbf{Z}^N(s-)) = \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))}(\mathbf{DF})(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))\right)e^{-\int_{0+}^{-\tau(\mathbf{Z}^N(s-))}(\mathbf{D}\boldsymbol{\mathfrak{F}})(\phi_s\mathbf{Z}^N(s-))\,ds}.$$

Theorem 1 shows that $\tau(\mathbf{Z}^N(s-)) \xrightarrow{\mathcal{D}} 0$; taking the limit as $N\to\infty$ on both sides gives

$$(\mathbf{D}\boldsymbol{\pi})(\boldsymbol{\pi}(\mathbf{Z}(s-))) = \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}(s-)))}(\mathbf{DF})(\boldsymbol{\pi}(\mathbf{Z}(s-)))\right), \tag{34}$$

where we have used Corollary 1 to obtain convergence of $\mathbf{Z}^N(s-)$ to $\boldsymbol{\pi}(\mathbf{Z}(s-))$.

First, by assumption, $N\mu_{ij}^N(\mathbf{x}) \to \theta_{ij}(\mathbf{x})$, which combined with some simple manipulations show that multiplying (31) by $(\partial_i\pi_i)(\mathbf{Z}^N(s-))$ and summing over i yield the first two terms in (11).

In a similar fashion, observing that (32) converges to

$$\sigma(\boldsymbol{\pi}(\mathbf{Z}(s-))),$$

we obtain the third line in the generator (11).

It remains to show that (33) vanishes. To that end, we observe that

$$(\mathbf{D}\pi_i)(\mathbf{Z}^N(s-)) \cdot F^N(\mathbf{Z}^N(s-)) = \mathbf{e}_i \cdot \left((\mathbf{D}\boldsymbol{\pi})(\mathbf{Z}^N(s-))\mathbf{F}^N(\mathbf{Z}^N(s-)\right),$$

while from (19), we have

$$(\mathbf{D}\boldsymbol{\pi})(\mathbf{Z}^N(s-))\mathbf{F}^N(\mathbf{Z}^N(s-) = \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))}(\mathbf{DF})(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))\right)$$
$$\times\, e^{-\int_{0+}^{-\tau(\mathbf{Z}^N(s-))}(\mathbf{D}\boldsymbol{\mathfrak{F}})(\phi_s\mathbf{Z}^N(s-))\,ds}\boldsymbol{\mathfrak{F}}(\mathbf{Z}^N(s-))R(\mathbf{Z}^N(s-)).$$

Now, (22) yields

$$e^{-\int_{0+}^{-\tau(\mathbf{Z}^N(s-))}(\mathbf{D}\boldsymbol{\mathfrak{F}})(\phi_s\mathbf{Z}^N(s-))\,ds}\boldsymbol{\mathfrak{F}}(\mathbf{Z}^N(s-)) = \boldsymbol{\mathfrak{F}}(\phi_{-\tau(\mathbf{Z}^N(s-))}\mathbf{Z}^N(s-)) = \boldsymbol{\mathfrak{F}}(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))$$

Thus, recalling from (8) that $\lambda(\mathbf{x}) = (\mathbf{D}R)(\mathbf{x}) \cdot \boldsymbol{\mathcal{F}}(\mathbf{x})$ we have

$$\left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))}(\mathbf{DF})(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))\right)\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{Z}(s-)))$$
$$= \left(\mathbf{I} - \frac{1}{\lambda(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))}\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))(\mathbf{D}R)(\boldsymbol{\pi}(\mathbf{Z}^N(s-)))\right)\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{Z}(s-)))$$
$$= \left(\boldsymbol{\mathcal{F}}(\boldsymbol{\pi}(\mathbf{Z}(s-))) - \frac{\boldsymbol{\mathcal{F}}(\mathbf{Z}(s-))\lambda(\boldsymbol{\pi}(\mathbf{Z}(s-))}{\lambda(\boldsymbol{\pi}(\mathbf{Z}(s-)))}\right) = \mathbf{0}.$$

9.5.2. *Quadratic variation terms.* We next consider the terms

$$\sum_{i=1}^{K}\int_{0+}^{t}(\partial_i^2\pi_i)(\mathbf{Z}^N(s-))d[M_i^N]_s.$$

We begin by recalling from (**??**) that

$$\left[M_i^N\right]_t = \frac{1}{N^2}\left(\sum_{\mathbf{n}\in\mathbb{N}_0^K} n_i^2 B_{i,\mathbf{n}}^N(Nt) + D_i^N(Nt)\right),$$

while rescaling and rearranging (1) gives

$$\frac{1}{N^2}B_{i,\mathbf{n}}^N(Nt) = \int_{0+}^{t}\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s))Z_i^N(s)\,ds + \frac{1}{N^2}\tilde{B}_{i,\mathbf{n}}^N(Nt).$$

Moreover,

$$\mathbb{E}\left[\left(\tfrac{1}{N^2}\tilde{B}_{i,\mathbf{n}}^N(Nt)\right)^2\right] = \frac{1}{N^4}\mathbb{E}\left[\left[\tilde{B}_{i,\mathbf{n}}^N\right]_{Nt}\right] = \frac{1}{N^4}\mathbb{E}\left[B_{i,\mathbf{n}}^N(t)\right] = \frac{1}{N^2}\int_{0+}^{t}\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s))Z_i^N(s)\,ds,$$

so

$$\frac{1}{N^2}B_{i,\mathbf{n}}^N(Nt) - \int_{0+}^{t}\beta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s))Z_i^N(s)\,ds \to 0$$

in probability as $N \to \infty$. Similarly, $\frac{1}{N^2}D_i^N(Nt) - \int_{0+}^{t}\delta_{i,\mathbf{n}}^N(\mathbf{Z}^N(s))Z_i^N(s)\,ds \to 0$.

Combining these with assumptions (3) and (4), it follows that as $N \to \infty$,

$$\left[M_i^N\right]_t \to \int_{0+}^{t}\check{\beta}_i(\boldsymbol{\pi}(\mathbf{Z}(s)) + \delta_i(\boldsymbol{\pi}(\mathbf{Z}(s)))\,ds. \tag{35}$$

Lastly, noting that for $\mathbf{x} \in \Omega$, $\delta_i(\mathbf{x}) = \bar{\beta}_i(\mathbf{x})$, we have

$$\mathbb{P} - \lim_{N\to\infty}\left[M_i^N\right]_t = \int_{0+}^{t}\check{\beta}_i(\boldsymbol{\pi}(\mathbf{Z}(s)) + \bar{\beta}_i(\boldsymbol{\pi}(\mathbf{Z}(s)))\,ds.$$

Lengthy, yet essentially rote, calculations using (**??**) and (35) show that in the limit as $N \to \infty$,

$$\sum_{i=1}^{K}\int_{0+}^{t}(\partial_i^2\pi_i)(\mathbf{Z}^N(s-))d[M_i^N]_s$$

gives the fourth through eighth lines in (11).

9.5.3. *Martingale terms.* Then,

$$\left[W^N_{i,n,b}\right]_t = \frac{1}{N^2}\int_{0+}^{t} \frac{d\left[\tilde{B}^N_{i,n\mathbf{e}_i}\right]_{Ns}}{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)}$$
$$= \frac{1}{N^2}\int_{0+}^{t} \frac{dB^N_{i,n\mathbf{e}_i}(Ns)}{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)},$$

which, using (1), equals

$$= \int_{0+}^{t} ds + \frac{1}{N}\int_{0+}^{t} \frac{d\tilde{B}^N_{i,n\mathbf{e}_i}(Ns)}{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)}.$$

Now,

$$\mathbb{E}\left[\left(\frac{1}{N}\int_{0+}^{t} \frac{d\tilde{B}^N_{i,n\mathbf{e}_i}(Ns)}{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)}\right)^2\right] = \frac{1}{N^2}\mathbb{E}\left[\int_{0+}^{t} \frac{d\left[\tilde{B}^N_{i,n\mathbf{e}_i}\right]_{Ns}}{\left(\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)\right)^2}\right]$$
$$= \frac{1}{N}\mathbb{E}\left[\int_{0+}^{t} \frac{ds}{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)}\right],$$

so $\left[W^N_{i,n,b}\right]_t \to t$ as $N \to \infty$. Since $W^N_{i,n,b}(t)$ is a martingale with jumps tending to 0 as $N \to \infty$, it follows from the Martingale central limit theorem that $W^N_{i,n,b} \xrightarrow{\mathcal{D}} W_{i,n,b}$, where the $W_{i,n,b}$ are independent Brownian motions. Similarly, the processes $W^N_{i,d}$ converge in distribution to Brownian motions $W_{i,d}$.

Rearranging, we observe that

$$\frac{1}{N}\tilde{B}^N_{i,n\mathbf{e}_i}(Nt) = \int_{0+}^{t} \sqrt{\beta^N_{i,n\mathbf{e}_i}(\mathbf{Z}^N(s))Z^N_i(s)}\, dW^N_{i,n,b}(s)$$

and

$$\frac{1}{N}\tilde{D}^N_i(Nt) = \int_{0+}^{t} \sqrt{\delta^N_i(\mathbf{Z}^N(s))Z^N_i(s)}\, dW^N_{i,d}(s).$$

Lastly, for $\mathbf{n} \neq |\mathbf{n}|\,\mathbf{e}_i$,

$$\frac{1}{N}\mathbb{E}\left[\left(\tilde{B}_{i,\mathbf{n}}(Nt)\right)^2\right] = \mathbb{E}\left[\int_{0+}^{t} \beta^N_{i,\mathbf{n}}(\mathbf{Z}^N(s))Z^N_i(s)\, ds\right],$$

which, by Assumption (5) is uniformly $O\left(\frac{1}{N}\right)$ on compact sets. Thus,

$$\frac{1}{N}\tilde{B}_{i,\mathbf{n}}(Nt) \xrightarrow{\mathcal{D}} 0.$$

9.5.4. *Remainder terms.* Lastly, from Taylor's Theorem, we have

$$\pi_k(\mathbf{x}+\Delta\mathbf{x}) - \pi_k(\mathbf{x}) - \sum_{i=1}^{K}(\partial_i\pi_k)(\mathbf{x})\Delta x_i(s) = \frac{1}{2}\sum_{1\le i,j\le K}\Delta x_i\Delta x_j\int_0^1 (1-t)^2(\partial_i\partial_j\pi_k)(\mathbf{x}+t\Delta\mathbf{x})\,dt$$

so that

$$\eta^N(t) = \sum_{0<s\le t} \left\{ \frac{1}{2} \sum_{1\le i,j\le K} \Delta Z_i^N(s) \Delta Z_j^N(s) \int_0^1 (1-t)^2 (\partial_i\partial_j\pi_k)(\mathbf{Z}^N(s-)+t\Delta\mathbf{Z}^N(s)) - (\partial_i\partial_j\pi_k)(\mathbf{Z}^N(s-))\, dt \right\}$$

and

$$\left|\eta^N(t)\right| \le \sum_{0<s\le t} \left\{ \frac{1}{2} \sum_{1\le i,j\le K} \Delta Z_i^N(s) \Delta Z_j^N(s) \sup_{\{\mathbf{h}:\|\mathbf{h}\|\le \frac{1}{N}\}} \left| (\partial_i\partial_j\pi_k)(\mathbf{Z}^N(s-)+\mathbf{h}) - (\partial_i\partial_j\pi_k)(\mathbf{Z}^N(s-)) \right| \right\}.$$

From our expression (**??**) for $(\partial_i\partial_j\pi_k)(\mathbf{x})$, it can be seen that this function is locally Lipschitz continuous, so

$$\left|\eta^N(t)\right| \ll \frac{1}{N}[\mathbf{Z}^N]_t = \frac{1}{N}[\mathbf{M}^N]_t,$$

so $\left|\eta^N(t)\right|$ weakly converges to 0 as $N \to \infty$, by arguments similar to those above.

9.5.5. *Well-posedness for the Martingale problem.*

9.6. **Proof of Corollary 2.** Let $\varepsilon_N = \frac{1-s}{4B_{\mathbf{x},\varepsilon}^2}\frac{\ln N}{N}$, where $0 < r' < r < s < 1$ and B are as in Propositions 1 and 2. Then, the aforementioned propositions tells us that

$$\mathbb{P}\{\sup_{t\le\varepsilon_N} \left|\mathbf{Z}^N(t) - \psi_{Nt}\mathbf{Z}^N(0)\right| > N^{-r}\} \le N^{-r},$$

while

$$\mathbb{P}\{\tau(\mathbf{Z}^N(\varepsilon_N) > N^{-r'}\} < N^{-r}.$$

In particular $\mathbf{Z}^N(\varepsilon_N) \xrightarrow{\mathcal{D}} \boldsymbol{\pi}(\mathbf{Z}(0)) \in \Omega$. Now, consider the process

$$\hat{\mathbf{Z}}^N(t) = \mathbf{Z}^N(t+\varepsilon_N).$$

By the above, $\hat{\mathbf{Z}}^N(0) \xrightarrow{\mathcal{D}} \hat{\mathbf{Z}}(0) \in \Omega$, while $\hat{\mathbf{Z}}^N$ satisfies (23). Thus, we may apply Theorem 2 to conclude that $\hat{\mathbf{Z}}^N \xrightarrow{\mathcal{D}} \hat{\mathbf{Z}}$, where the limiting process has generator (11). Moreover,

$$\tilde{\mathbf{Z}}^N(t+\varepsilon_N) - \hat{\mathbf{Z}}^N(t) = \psi_{Nt+N\varepsilon_N}\mathbf{Z}^N(0) - \boldsymbol{\pi}(\mathbf{Z}^N(0)).$$

Since $N\varepsilon_N \to \infty$ as $N \to \infty$, the latter converges to 0 uniformly in t, so we must have $\tilde{\mathbf{Z}}^N(\cdot + \varepsilon_N) - \hat{\mathbf{Z}}^N \xrightarrow{\mathcal{D}} 0$.

Lastly, define $\gamma^N : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$\gamma^N(t) \stackrel{\text{def}}{=} \begin{cases} 0 & 0 \le \text{if } t \le \varepsilon_N \\ t - \varepsilon_N & \text{otherwise.} \end{cases}$$

Then, $\gamma^N(t) \to t$ uniformly in N, while

$$\hat{\mathbf{Z}}^N(\gamma^N(t)) = \begin{cases} \hat{\mathbf{Z}}^N(0) & 0 \le \text{if } t \le \varepsilon_N \\ \tilde{\mathbf{Z}}^N(t) & \text{otherwise.} \end{cases}$$

Now, the map $(f,g) \to f \circ g : \mathbb{D}_\Omega[0,\infty) \times \mathbb{D}_{\mathbb{R}+}[0,\infty) \to \mathbb{D}_\Omega[0,\infty)$ is continuous at points where the pair f,g are continuous. Since $\hat{\mathbf{Z}}$ is continuous, we may apply the continuous mapping theorem to conclude that $\hat{\mathbf{Z}}^N \circ \gamma^N \xrightarrow{\mathcal{D}} \hat{\mathbf{Z}}$ as $N \to \infty$; on the other hand,

$$\sup_t \left|\tilde{\mathbf{Z}}^N(t) - \hat{\mathbf{Z}}^N(\gamma^N(t))\right| = \sup_{t \leq \varepsilon_N} \left|\tilde{\mathbf{Z}}^N(t) - \tilde{\mathbf{Z}}^N(\varepsilon_N)\right|$$

which converges uniformly to 0. Thus $\tilde{\mathbf{Z}}^N - \hat{\mathbf{Z}}^N \xrightarrow{\mathcal{D}} 0$, and the result follows.

References

[1] R. E. Bellman, *Introduction to matrix analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 1995.

[2] P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.

[3] J. H. Gillespie, *Nautural selection for within-generation variance in offspring number.*, Genetics **76** (1974), no. 3, 601–606.

[4] ———, *Natural selection for within-generation variance in offspring number ii. discrete haploid models.*, Genetics **81** (1975), no. 2, 403–413.

[5] M. W. Hirsch, *Systems of differential equations which are competitive or cooperative: III. Competing species*, Nonlinearity **1** (1988), 51–71.

[6] M. W. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra*, Academic Press, New York, 1974.

[7] T. G. Kurtz and P. Protter, *Weak limit theorems for stochastic integrals and stochastic differential equations*, Ann. Probab. **19** (1991), no. 3, 1035–1070.

[8] ———, *Weak limit theorems for stochastic integrals and stochastic differential equations II: Infinite dimensional case*, Probabilistic Models for Nonlinear Partial Differential Equations, Lecture Notes in Mathematics, vol. 1627, Springer, 1996, pp. 197–285.

[9] T. L. Parsons, C. Quince, and J. B. Plotkin, *Some consequences of demographic stochasticity in population genetics*, in preparation, 2010.

[10] P. E. Protter, *Stochastic integration and differential equations*, 2nd ed., Springer, Berlin, 2004.